

Serial Number: 09/938,901**ENTERED**

- ☐ Changed a file from non-ASCII to ASCII
- ☐ Changed the margins in cases where the sequence text was "wrapped" down to the next line.
- ☐ Edited a format error in the Current Application Data section, specifically: _____
- ☐ Edited the Current Application Data section with the actual current number. The number inputted by the applicant was ☐ the prior application data; or ☐ other _____
- ☐ Added the mandatory heading and subheadings for "Current Application Data".
- ☐ Edited the "Number of Sequences" field. The applicant spelled out a number instead of using an integer.
- ☐ Changed the spelling of a mandatory field (the headings or subheadings), specifically: _____
- ☐ Corrected the SEQ ID NO when obviously incorrect. The sequence numbers that were edited were: _____
- ☐ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited: _____
- ☐ Corrected subheading placement. All responses must be on the same line as each subheading. If the applicant placed a response below the subheading, this was moved to its appropriate place.
- ☐ Inserted colons after headings/subheadings. Headings edited included: _____
- ☐ Deleted extra, invalid, headings used by an applicant, specifically: _____
- ☒ Deleted: ☒ non-ASCII "garbage" at the beginning/end of files; ☐ secretary initials/filename at end of file.
☐ page numbers throughout text; ☐ other invalid text, such as _____
- ☐ Inserted mandatory headings, specifically: _____
- ☐ Corrected an obvious error in the response, specifically: _____
- ☐ Edited identifiers where upper case is used but lower case is required, or vice versa.
- ☐ Corrected an error in the Number of Sequences field, specifically: _____
- ☐ A "Hard Page Break" code was inserted by the applicant. All occurrences had to be deleted.
- ☐ Deleted *ending* stop codon in amino acid sequences and adjusted the "(A)Length:" field accordingly (error due to a PatentIn bug). Sequences corrected: _____
- ☐ Other: _____

Examiner: The above corrections must be communicated to the applicant in the first Office Action. DO NOT send a copy of this form

2/1/95

RAW SEQUENCE LISTING

DATE: 10/16/2001

PATENT APPLICATION: US/09/938,901

TIME: 17:37:13

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF3\10162001\I938901.raw

```

3 <110> APPLICANT: Kuramitsu Seiki,
4   Yokoyama Shigeyuki
6 <120> TITLE OF INVENTION: GENE ENCODING DNA REPAIR ENZYME
8 <130> FILE REFERENCE: PH-1261-US
C--> 10 <140> CURRENT APPLICATION NUMBER: US/09/938,901
C--> 11 <141> CURRENT FILING DATE: 2001-08-24
13 <150> PRIOR APPLICATION NUMBER: JP2001-47762
14 <151> PRIOR FILING DATE: 2001-02-23
16 <160> NUMBER OF SEQ ID NOS: 17
18 <170> SOFTWARE: PatentIn Ver. 2.0
20 <210> SEQ ID NO: 1
21 <211> LENGTH: 975
22 <212> TYPE: DNA
23 <213> ORGANISM: Thermus thermophilus
25 <220> FEATURE:
26 <221> NAME/KEY: CDS
27 <222> LOCATION: (1)..(975)
29 <400> SEQUENCE: 1
30 gtg gag gcc tgg cgg aaa gcc ctc ctc gcc tgg tac cgg gaa aac gcc 48
31 Val Glu Ala Trp Arg Lys Ala Leu Leu Ala Trp Tyr Arg Glu Asn Ala
32 1 5 10 15
34 cgc ccc ctc ccc tgg cgg ggg gag aag gac cct tac cgc gtc ctg gtc 96
35 Arg Pro Leu Pro Trp Arg Gly Glu Lys Asp Pro Tyr Arg Val Leu Val
36 20 25 30
38 tcc gag gtc ctt ctg cag cag acc cgg gtg gag cag gcc ctc ccc tat 144
39 Ser Glu Val Leu Leu Gln Gln Thr Arg Val Glu Gln Ala Leu Pro Tyr
40 35 40 45
42 tac cgc cgc ttt ctg gag cgc ttt ccc acc ctg aag gcc ctg gcc gcg 192
43 Tyr Arg Arg Phe Leu Glu Arg Phe Pro Thr Leu Lys Ala Leu Ala Ala
44 50 55 60
46 gct tcc ctg gaa gag gtc ctt agg gtc tgg cag ggg gcg ggc tac tac 240
47 Ala Ser Leu Glu Glu Val Leu Arg Val Trp Gln Gly Ala Gly Tyr Tyr
48 65 70 75 80
50 cgg cgg gcg gaa cac ctc cac cgc ctg gcc cga agc gtg gag gag ctt 288
51 Arg Arg Ala Glu His Leu His Arg Leu Ala Arg Ser Val Glu Glu Leu
52 85 90 95
54 ccc ccg agc ttc gcc gag ctt cgg ggg ctt cct ggt ctc ggg cct tac 336
55 Pro Pro Ser Phe Ala Glu Leu Arg Gly Leu Pro Gly Leu Gly Pro Tyr
56 100 105 110
58 acc gcg gcg gcg gtg gcc tcc atc gcc ttc ggg gag cgg gtg gcg gcg 384
59 Thr Ala Ala Ala Val Ala Ser Ile Ala Phe Gly Glu Arg Val Ala Ala
60 115 120 125
62 gtg gac ggg aac gtc cgg agg gtc ctc tcc cgc ctc ttc gcc cgg gaa 432
63 Val Asp Gly Asn Val Arg Arg Val Leu Ser Arg Leu Phe Ala Arg Glu
64 130 135 140
66 agc ccc aag gag aag gag ctt ttc gcc ctc gcc cag ggc ctc ctc ccc 480
67 Ser Pro Lys Glu Lys Glu Leu Phe Ala Leu Ala Gln Gly Leu Leu Pro

```

RAW SEQUENCE LISTING

DATE: 10/16/2001

PATENT APPLICATION: US/09/938,901

TIME: 17:37:13

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF3\10162001\I938901.raw

```

68 145          150          155          160
70 gag ggc gtg gac ccg ggg gtg tgg aac cag gcc ctc atg gag ctc ggg 528
71 Glu Gly Val Asp Pro Gly Val Trp Asn Gln Ala Leu Met Glu Leu Gly
72          165          170          175
74 gcc acg gtc tgc ctg ccg aaa cgg ccc cgt tgc ggg gcc tgc ccc cta 576
75 Ala Thr Val Cys Leu Pro Lys Arg Pro Arg Cys Gly Ala Cys Pro Leu
76          180          185          190
78 ggg gcc ttc tgc cgg ggg aag gag gcc ccc ggg cgc tac ccc gcg ccc 624
79 Gly Ala Phe Cys Arg Gly Lys Glu Ala Pro Gly Arg Tyr Pro Ala Pro
80          195          200          205
82 agg aag cgc cgg gcg aag gag gag cgc ctc gtc gcc ctc gtc ctc ctc 672
83 Arg Lys Arg Arg Ala Lys Glu Glu Arg Leu Val Ala Leu Val Leu Leu
84          210          215          220
86 ggg cgg aag ggg gtg cac ctg gaa agg ctt gag ggg cgc ttc cag ggc 720
87 Gly Arg Lys Gly Val His Leu Glu Arg Leu Glu Gly Arg Phe Gln Gly
88 225          230          235          240
90 ctc tac ggc gtc ccc ctc ttt ccc cct gag gag ctt ccc ggg cgg gag 768
91 Leu Tyr Gly Val Pro Leu Phe Pro Pro Glu Glu Leu Pro Gly Arg Glu
92          245          250          255
94 gcg gcc ttc ggg gtg agg tct agg ccc cta ggc gag gtg cgc cac gcc 816
95 Ala Ala Phe Gly Val Arg Ser Arg Pro Leu Gly Glu Val Arg His Ala
96          260          265          270
98 ctc acc cac cgg agg ctt cgc gtg gag gtg cgg ggg gcc ctt tgg gaa 864
99 Leu Thr His Arg Arg Leu Arg Val Glu Val Arg Gly Ala Leu Trp Glu
100          275          280          285
102 ggg gag ggg gag gac ccc tgg aag agg ccc cta ccc aag ctc atg gag 912
103 Gly Glu Gly Glu Asp Pro Trp Lys Arg Pro Leu Pro Lys Leu Met Glu
104          290          295          300
106 aag gtg ctc cgc aag gcg ctt ccc ctc ctc gct cat gcg ggc gta gtc 960
107 Lys Val Leu Arg Lys Ala Leu Pro Leu Leu Ala His Ala Gly Val Val
108 305          310          315          320
110 ccc ctc ccg gac gca 975
111 Pro Leu Pro Asp Ala
112          325
115 <210> SEQ ID NO: 2
116 <211> LENGTH: 325
117 <212> TYPE: PRT
118 <213> ORGANISM: Thermus thermophilus
120 <400> SEQUENCE: 2
121 Val Glu Ala Trp Arg Lys Ala Leu Leu Ala Trp Tyr Arg Glu Asn Ala
122 1          5          10          15
124 Arg Pro Leu Pro Trp Arg Gly Glu Lys Asp Pro Tyr Arg Val Leu Val
125          20          25          30
127 Ser Glu Val Leu Leu Gln Gln Thr Arg Val Glu Gln Ala Leu Pro Tyr
128          35          40          45
130 Tyr Arg Arg Phe Leu Glu Arg Phe Pro Thr Leu Lys Ala Leu Ala Ala
131          50          55          60
133 Ala Ser Leu Glu Glu Val Leu Arg Val Trp Gln Gly Ala Gly Tyr Tyr
134 65          70          75          80

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/938,901

DATE: 10/16/2001

TIME: 17:37:13

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF3\10162001\I938901.raw

```

136 Arg Arg Ala Glu His Leu His Arg Leu Ala Arg Ser Val Glu Glu Leu
137                               85                               90                               95
139 Pro Pro Ser Phe Ala Glu Leu Arg Gly Leu Pro Gly Leu Gly Pro Tyr
140                               100                               105                               110
142 Thr Ala Ala Val Ala Ser Ile Ala Phe Gly Glu Arg Val Ala Ala
143                               115                               120                               125
145 Val Asp Gly Asn Val Arg Arg Val Leu Ser Arg Leu Phe Ala Arg Glu
146                               130                               135                               140
148 Ser Pro Lys Glu Lys Glu Leu Phe Ala Leu Ala Gln Gly Leu Leu Pro
149 145                               150                               155                               160
151 Glu Gly Val Asp Pro Gly Val Trp Asn Gln Ala Leu Met Glu Leu Gly
152                               165                               170                               175
154 Ala Thr Val Cys Leu Pro Lys Arg Pro Arg Cys Gly Ala Cys Pro Leu
155                               180                               185                               190
157 Gly Ala Phe Cys Arg Gly Lys Glu Ala Pro Gly Arg Tyr Pro Ala Pro
158                               195                               200                               205
160 Arg Lys Arg Arg Ala Lys Glu Glu Arg Leu Val Ala Leu Val Leu Leu
161                               210                               215                               220
163 Gly Arg Lys Gly Val His Leu Glu Arg Leu Glu Gly Arg Phe Gln Gly
164 225                               230                               235                               240
166 Leu Tyr Gly Val Pro Leu Phe Pro Pro Glu Glu Leu Pro Gly Arg Glu
167                               245                               250                               255
169 Ala Ala Phe Gly Val Arg Ser Arg Pro Leu Gly Glu Val Arg His Ala
170                               260                               265                               270
172 Leu Thr His Arg Arg Leu Arg Val Glu Val Arg Gly Ala Leu Trp Glu
173                               275                               280                               285
175 Gly Glu Gly Glu Asp Pro Trp Lys Arg Pro Leu Pro Lys Leu Met Glu
176                               290                               295                               300
178 Lys Val Leu Arg Lys Ala Leu Pro Leu Leu Ala His Ala Gly Val Val
179 305                               310                               315                               320
181 Pro Leu Pro Asp Ala
182                               325
185 <210> SEQ ID NO: 3
186 <211> LENGTH: 1998
187 <212> TYPE: DNA
188 <213> ORGANISM: Thermus thermophilus
190 <220> FEATURE:
191 <221> NAME/KEY: CDS
192 <222> LOCATION: (1)..(1998)
194 <400> SEQUENCE: 3
195 atg agg gac cgg gtc cgc tgg cgg gtg ctt tcc ctc cct ccc ctc gcc 48
196 Met Arg Asp Arg Val Arg Trp Arg Val Leu Ser Leu Pro Pro Leu Ala
197 1 5 10 15
199 cag tgg cgg gag gtg atg gcg gcc ttg gag gtg ggg ccg gag gcc gcc 96
200 Gln Trp Arg Glu Val Met Ala Ala Leu Glu Val Gly Pro Glu Ala Ala
201 20 25 30
203 ctg gcc tac tgg cac cgg ggc ttt agg cgc aag gag gac ctg gac ccc 144
204 Leu Ala Tyr Trp His Arg Gly Phe Arg Arg Lys Glu Asp Leu Asp Pro
205 35 40 45

```

RAW SEQUENCE LISTING

DATE: 10/16/2001

PATENT APPLICATION: US/09/938,901

TIME: 17:37:13

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF3\10162001\I938901.raw

```

207 ccc ctc gcc ctc ctt ccc ctc aag ggc ctg agg gag gcg gcg gcc ctc 192
208 Pro Leu Ala Leu Leu Pro Leu Lys Gly Leu Arg Glu Ala Ala Ala Leu
209      50      55      60
211 ctg gag gag gcg ctc cgc cag ggg aag cgg atc cgc gtc cac ggg gac 240
212 Leu Glu Glu Ala Leu Arg Gln Gly Lys Arg Ile Arg Val His Gly Asp
213 65      70      75      80
215 tac gac gcc gac ggg ctc acg ggc acg gcc atc ctg gtt cgg ggc ctc 288
216 Tyr Asp Ala Asp Gly Leu Thr Gly Thr Ala Ile Leu Val Arg Gly Leu
217      85      90      95
219 gcc gcc ttg ggc gcc gac gtc cac ccc ttc atc ccc cac cgg ctg gag 336
220 Ala Ala Leu Gly Ala Asp Val His Pro Phe Ile Pro His Arg Leu Glu
221      100      105      110
223 gaa ggg tac ggg gtg ctg atg gag cgg gtt ccc gag cac ctc gag gcc 384
224 Glu Gly Tyr Gly Val Leu Met Glu Arg Val Pro Glu His Leu Glu Ala
225      115      120      125
227 tcg gac ctc ttc ctc acc gtg gac tgc ggg atc acg aac cac gcc gag 432
228 Ser Asp Leu Phe Leu Thr Val Asp Cys Gly Ile Thr Asn His Ala Glu
229      130      135      140
231 ctc agg gag ctt ttg gaa aac ggg gtg gag gtg atc gtc acc gac cac 480
232 Leu Arg Glu Leu Leu Glu Asn Gly Val Glu Val Ile Val Thr Asp His
233 145      150      155      160
235 cac acc ccc ggc aag acc cct tcc ccc ggc ctc gtg gtc cac ccc gcc 528
236 His Thr Pro Gly Lys Thr Pro Ser Pro Gly Leu Val Val His Pro Ala
237      165      170      175
239 ctc acc ccg gac ctt aag gag aag ccc acg ggg gcg ggg gtg gtc ttc 576
240 Leu Thr Pro Asp Leu Lys Glu Lys Pro Thr Gly Ala Gly Val Val Phe
241      180      185      190
243 ctc ctc ctc tgg gcc ctc cac gag cgc ctg ggc ctt ccc cca ccc ctg 624
244 Leu Leu Leu Trp Ala Leu His Glu Arg Leu Gly Leu Pro Pro Pro Leu
245      195      200      205
247 gag tac gcc gac ctc gcc gcg gtg ggc acc atc gcc gac gtg gcc ccc 672
248 Glu Tyr Ala Asp Leu Ala Ala Val Gly Thr Ile Ala Asp Val Ala Pro
249      210      215      220
251 ctt tgg ggc tgg aac cgg gcc ttg gtg aag gag ggc ctg gcc cgc atc 720
252 Leu Trp Gly Trp Asn Arg Ala Leu Val Lys Glu Gly Leu Ala Arg Ile
253 225      230      235      240
255 ccc gcc tcc tcc tgg gtt ggg ctc agg ctt ctg gcc gag gcg gtg ggg 768
256 Pro Ala Ser Ser Trp Val Gly Leu Arg Leu Leu Ala Glu Ala Val Gly
257      245      250      255
259 tac acg ggg aag gcg gtg gag gtg gcc ttc cgc atc gcc ccc cgg atc 816
260 Tyr Thr Gly Lys Ala Val Glu Val Ala Phe Arg Ile Ala Pro Arg Ile
261      260      265      270
263 aac gcg gca agc cgc ctc ggg gag gct gag aag gcc cta agg ctc ctc 864
264 Asn Ala Ala Ser Arg Leu Gly Glu Ala Glu Lys Ala Leu Arg Leu Leu
265      275      280      285
267 ctc acc gac gac gcg gcc gag gcc cag gcc ctc gtg ggg gaa ctc cac 912
268 Leu Thr Asp Asp Ala Ala Glu Ala Gln Ala Leu Val Gly Glu Leu His
269      290      295      300
271 cgg ctg aac gcc cgc cgc cag acc ctg gag gag gcc atg ctc agg aag 960

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/938,901

DATE: 10/16/2001

TIME: 17:37:13

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF3\10162001\I938901.raw

272	Arg	Leu	Asn	Ala	Arg	Arg	Gln	Thr	Leu	Glu	Glu	Ala	Met	Leu	Arg	Lys	
273	305						310					315				320	
275	ctc	ctc	ccc	cag	gcg	gac	ccc	gag	gcc	aag	gcc	atc	gtc	ctc	ctg	gac	1008
276	Leu	Leu	Pro	Gln	Ala	Asp	Pro	Glu	Ala	Lys	Ala	Ile	Val	Leu	Leu	Asp	
277				325						330						335	
279	ccc	gag	ggg	cac	ccg	ggg	gtg	atg	ggc	atc	gtg	gcg	agc	cgc	atc	ctg	1056
280	Pro	Glu	Gly	His	Pro	Gly	Val	Met	Gly	Ile	Val	Ala	Ser	Arg	Ile	Leu	
281				340						345						350	
283	gag	gcc	acc	ctc	cgg	ccc	gtc	ttc	ctg	gtg	gcc	cag	ggc	aag	ggg	acg	1104
284	Glu	Ala	Thr	Leu	Arg	Pro	Val	Phe	Leu	Val	Ala	Gln	Gly	Lys	Gly	Thr	
285			355					360						365			
287	gtg	cgg	agc	ctc	gcc	ccc	atc	agc	gcc	gtg	gag	gcc	cta	agg	agc	gcc	1152
288	Val	Arg	Ser	Leu	Ala	Pro	Ile	Ser	Ala	Val	Glu	Ala	Leu	Arg	Ser	Ala	
289		370					375						380				
291	gag	gac	ctt	ttg	ttg	cgc	tac	ggg	ggg	cac	aag	gag	gcg	gcg	ggc	ttc	1200
292	Glu	Asp	Leu	Leu	Leu	Arg	Tyr	Gly	Gly	His	Lys	Glu	Ala	Ala	Gly	Phe	
293	385					390					395					400	
295	gcc	atg	gac	gag	gcc	ctc	ttc	ccc	gcc	ttc	aag	gcc	cgg	gtg	gag	gcc	1248
296	Ala	Met	Asp	Glu	Ala	Leu	Phe	Pro	Ala	Phe	Lys	Ala	Arg	Val	Glu	Ala	
297				405						410						415	
299	tac	gcc	gcc	cgc	ttc	ccc	gac	ccc	gtg	cgc	gag	gtg	gcc	ctt	ttg	gac	1296
300	Tyr	Ala	Ala	Arg	Phe	Pro	Asp	Pro	Val	Arg	Glu	Val	Ala	Leu	Leu	Asp	
301				420					425						430		
303	ctg	ctt	ccg	gag	ccc	ggc	ctc	ctc	ccc	cag	gtc	ttc	cgg	gag	ctc	gcc	1344
304	Leu	Leu	Pro	Glu	Pro	Gly	Leu	Leu	Pro	Gln	Val	Phe	Arg	Glu	Leu	Ala	
305			435					440						445			
307	ctt	ttg	gag	ccc	tac	ggc	gag	gga	aac	ccc	gag	ccc	ctc	ttc	ctc	ctc	1392
308	Leu	Leu	Glu	Pro	Tyr	Gly	Glu	Gly	Asn	Pro	Glu	Pro	Leu	Phe	Leu	Leu	
309		450					455					460					
311	ttc	ggc	gcc	ccg	gag	gag	gcc	cgg	cgc	ctc	ggg	gag	ggc	cgc	cac	ctc	1440
312	Phe	Gly	Ala	Pro	Glu	Glu	Ala	Arg	Arg	Leu	Gly	Glu	Gly	Arg	His	Leu	
313	465					470					475					480	
315	gcc	ttc	cgc	ctg	aag	ggg	gtg	cgg	gtc	ctg	gcc	tgg	aaa	cag	ggg	gac	1488
316	Ala	Phe	Arg	Leu	Lys	Gly	Val	Arg	Val	Leu	Ala	Trp	Lys	Gln	Gly	Asp	
317				485						490						495	
319	ctc	gcc	ctg	ccc	ccg	gag	gtg	gag	gtg	gcg	ggc	ctc	ctc	agc	gaa	aac	1536
320	Leu	Ala	Leu	Pro	Pro	Glu	Val	Glu	Val	Ala	Gly	Leu	Leu	Ser	Glu	Asn	
321				500						505					510		
323	gcc	tgg	aac	ggc	cac	ctc	gcc	tac	gag	gtc	cag	gcg	gtg	gac	ctg	cga	1584
324	Ala	Trp	Asn	Gly	His	Leu	Ala	Tyr	Glu	Val	Gln	Ala	Val	Asp	Leu	Arg	
325			515					520						525			
327	aag	cca	gag	gcg	ctg	gag	ggc	ggg	atc	gcg	ccc	ttc	gcc	tac	ccc	ctg	1632
328	Lys	Pro	Glu	Ala	Leu	Glu	Gly	Gly	Ile	Ala	Pro	Phe	Ala	Tyr	Pro	Leu	
329		530					535					540					
331	ccc	ctc	ctc	gag	gcc	ctg	gcc	cgg	gcc	cgc	ctg	ggg	gaa	ggg	gtc	tac	1680
332	Pro	Leu	Leu	Glu	Ala	Leu	Ala	Arg	Ala	Arg	Leu	Gly	Glu	Gly	Val	Tyr	
333	545					550					555					560	
335	gtc	ccc	gag	gac	aac	cct	gag	ggg	ctg	gac	tac	gcc	agg	aag	gcg	ggc	1728
336	Val	Pro	Glu	Asp	Asn	Pro	Glu	Gly	Leu	Asp	Tyr	Ala	Arg	Lys	Ala	Gly	

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/938,901

DATE: 10/16/2001

TIME: 17:37:14

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF3\10162001\I938901.raw

L:10 M:270 C: Current Application Number differs, Replaced Application Number

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date